2) Schumacher's Source Coding Theorem.

We will distinguish the cases of a source emitting pure states $|\varphi_x\rangle$ with probability p_x and that of a source emitting mixed states ρ_x with probability p_x.

2.a) Source of pure states.

Consider a source emitting states $|\varphi_x\rangle$ with probability p_x.
A message is a string of N "letters":

$$|\varphi_{x_1}\rangle \otimes \cdots \otimes |\varphi_{x_N}\rangle$$

and has probability $p_{x_1} p_{x_2} \cdots p_{x_N}$. If the states $|\varphi_x\rangle$ are mutually orthogonal we are in the same situation than in the classical Shannon Theory. Words are strings $x_1 \ldots x_N$ with probabilities $p_{x_1} p_{x_2} \cdots p_{x_N}$. The number of "typical words" is given by $H(p_{x_1} \cdots p_{x_N}) = N H(X)$. So we expect that it is possible to faithfully describe this source with NR bits as long as $R > H(X)$. On the other hand it is not possible to faithfully describe the source if $R < H(X)$.

We will see that in the quantum case the same theorem holds with $H(X)$ replaced by $S(\rho)$. Since in general $S(\rho) < H(X)$ quantum mechanically it is possible to compress more than in the classical case. This is essentially because the states $\{|\varphi_x\rangle\}$ if not - mutually orthogonal are redundant (they cannot be perfectly distinguished).

The method of proof uses as in the classical case the idea of typicality.

Typicality in the classical case.

Let X be a random variable with $\mathrm{Prob}\,(X = x) = p_x$.

N- strings are said to be ε-typical if they belong to the set

$$A_{N,\varepsilon}(X) = \Big\{ (x_1 \ldots x_N) \text{ is a string such that } \Big| -\frac{1}{N} \sum_{i=1}^{N} \log_2 p(x_i) - H(X) \Big| < \varepsilon \Big\}.$$

The importance of this definition comes from the following

Theorem on typical sequences.

Fix any $\varepsilon > 0$, $\delta > 0$ (small). We can find N sufficiently large such that:

(1) $\mathrm{Prob}\,(A_{N,\varepsilon}(X)) \geq 1 - \delta$.

(2) $(1-\delta)\, 2^{N(H(X)-\varepsilon)} \leq |A_{N,\varepsilon}(X)| \leq 2^{N(H(X)+\varepsilon)}$

(3) Let S_N a collection of strings $(x_1 \ldots x_N)$ with $|S_N| \leq 2^{NR}$ with $R < H(X) - \varepsilon$. Then $\mathrm{Prob}(S_N) \leq \delta$.

Proof. Statement (1) is just the law of large numbers because $H(X) = \mathbb{E}_X \left(-\log p(x) \right)$ and $-\log p(x_i)$ are i.i.d.

Statement (2) follows from

$$1-\delta \leq \mathrm{Prob}\,(A_{N,\varepsilon}(X)) \leq 1$$

$$\left.\begin{aligned} & 1-\delta \leq \sum_{x_1 \dots x_N \in A_{N,\varepsilon}(X)} p_{x_1 \dots x_N} \leq 1 \\ \text{and}\quad & 2^{-N(H(X)-\varepsilon)} \leq p_{x_1 \dots x_N} \leq 2^{-N(H(X)+\varepsilon)} \end{aligned}\right\} \longrightarrow (2).$$

The third statement follows by splitting $S_N = \left(S_N \cap A^c_{N,\varepsilon}(X)\right) \cup \left(S_N \cap A_{N,\varepsilon}(X)\right)$.

$$\mathrm{Prob}\,(S_N \cap A^c_{N,\varepsilon}(X)) \leq \delta.$$

$$\begin{aligned} \mathrm{Prob}\,(S_N \cap A_{N,\varepsilon}(X)) &\leq 2^{NR} \cdot 2^{-N(H(X)-\varepsilon)} \\ &= 2^{-N(H(X)-R-\varepsilon)}. \\ &\to 0 \quad \text{if} \quad R < H(X)-\varepsilon. \end{aligned}$$

Typicality in the quantum case.

In the quantum case we have an analogous Theorem and definition of typicality. Let ρ be a given density matrix. This has a spectral decomposition

$$\rho = \sum_a \lambda_a |a\rangle\langle a| \quad \text{or} \quad \sum_a \lambda_a \underset{\substack{\uparrow \\ \text{possibly degenerate.}}}{P_a}$$

Here $\lambda_a \geq 0$; $\sum_a \lambda_a = 1$

$P_a = P_a^2$ and $\sum_a P_a = 1$. The Hilbert space is $\mathcal{H}$.

Consider the Hilbert space $\underbrace{\mathcal{H} \otimes \mathcal{H} \cdots \otimes \mathcal{H}}_{N \text{ times}} = \mathcal{H}^{\otimes N}$.

The projector $P_{N,\varepsilon}(\rho)$ onto the "typical subspace" of $\mathcal{H}^{\otimes N}$ is defined as

$$P_{N,\varepsilon}(\rho) = \sum_{\substack{a_1 \cdots a_N \\ \text{typical sequences}}} P_{a_1} \otimes P_{a_2} \cdots \otimes P_{a_N}$$

where

$$\text{typical sequences} = \left\{ (a_1 \ldots a_N) \;\middle|\; \left| -\frac{1}{N} \sum_{i=1}^{N} \log \lambda_{a_i} - S(\rho) \right| < \varepsilon \right\}.$$

We have the following theorem (Below $\rho^{\otimes N} = \underbrace{\rho \otimes \rho \otimes \cdots \otimes \rho}_{N \text{ times}}$).

<u>Theorem on the typical subspace.</u>

Fix any $\varepsilon > 0$, $\delta > 0$ (small). We can find N sufficiently large so that:

(1) $\mathrm{Tr}\left(P_{N,\varepsilon}(\rho) \cdot \rho^{\otimes N}\right) \geq 1 - \delta$

(2) $(1-\delta)\, 2^{N(S(\rho) - \varepsilon)} \leq \mathrm{Tr}\, P_{N,\varepsilon}(\rho) \leq 2^{N(S(\rho)+\varepsilon)}$

(3) Let S_N be a projector (i.e. $S_N^2 = S_N$) such that $\mathrm{Tr}\, S_N \leq 2^{NR}$ with $R < S(\rho) - \varepsilon$. Then

$$\mathrm{Tr}\left(S_N\, \rho^{\otimes N}\right) \leq \delta .$$

Proof.

Statement 1 : Observe that

$$\rho^{\otimes N} = \sum_{a_1 \cdots a_N} \lambda_{a_1} \lambda_{a_2} \cdots \lambda_{a_N} \underbrace{P_{a_1} \otimes P_{a_2} \otimes \cdots \otimes P_{a_N}}_{\text{mutually orthonormal projectors.}} .$$

$$\mathrm{Tr}\left(P_{N,\varepsilon}(\rho)\ \rho^{\otimes N}\right) = \mathrm{Tr} \sum_{\substack{a_1 \cdots a_N \\ \text{typical}}} \lambda_{a_1} \cdots \lambda_{a_N}\ P_{a_1} \otimes P_{a_2} \cdots \otimes P_{a_N}$$

$$= \sum_{\substack{a_1 \cdots a_N \\ \text{typical}}} \lambda_{a_1} \cdots \lambda_{a_N}$$

= probability of set of typical sequences where

$$\text{typical sequences} = \left\{ (a_1 \cdots a_N) \ \middle|\ \left| -\frac{1}{N}\sum_{i=1}^{N} \log \lambda_{a_i} - S(\rho) \right| < \varepsilon \right\} .$$

By the law of large numbers : for N sufficiently large :

$$\mathrm{Prob} \left\{ (a_1 \cdots a_N) \ \middle|\ \left| -\frac{1}{N}\sum_{i=1}^{N} \log \lambda_{a_i} - \underbrace{\sum_a \lambda_a(-\log \lambda_a)}_{S(\rho)} \right| < \varepsilon \right\} \geq 1 - \delta .$$

Therefore $\qquad \mathrm{Tr}\left(P_{N,\varepsilon}(\rho)\ \rho^{\otimes N}\right) \geqslant 1 - \delta .$

Statement 2 : Observe that $\mathrm{Tr}\ P_{N,\varepsilon}(\rho) = \sum_{\substack{a_1 \cdots a_N \\ \text{typical}}} 1 = \#(\text{of typ sequences}).$

So the argument is the same than in the classical case.

Statement 3 : Observe that :

$$\mathrm{Tr}\, S_N\, \rho^{\otimes N} = \mathrm{Tr}\left(S_N\, \rho^{\otimes N}\, P_{N,\varepsilon}(\rho)\right) + \mathrm{Tr}\, S_N\, \rho^{\otimes N}(I - P_{N,\varepsilon}(\rho))$$

$*$ $\mathrm{Tr}\, S_N\, \rho^{\otimes N} P_{N,\varepsilon}(\rho) = \mathrm{Tr}\, S_N\, \rho^{\otimes N} P_{N,\varepsilon}(\rho)\, S_N$ (cyclicity)

$$= \mathrm{Tr}\, S_N\, P_{N,\varepsilon}(\rho)\, \rho^{\otimes N}\, P_{N,\varepsilon}(\rho)\, S_N$$

($P_{N,\varepsilon}(\rho)$ and $\rho^{\otimes N}$ commute).

$$\leq 2^{-N(S(\rho)-\varepsilon)}\, \mathrm{Tr}\, S_N$$

$$\leq 2^{-N(S(\rho) - R - \varepsilon)} \longrightarrow 0 \quad \text{if} \quad R < S(\rho) - \varepsilon.$$

$*$ $\mathrm{Tr}\, S_N\, \rho^{\otimes N}(I - P_{N,\varepsilon}(\rho)) = \mathrm{Tr}\, S_N\, \rho^{\otimes N}(I - P_{N,\varepsilon}(\rho))$

$(I - P_{N,\varepsilon}(\rho))$

and $\underbrace{(I - P_{N,\varepsilon}(\rho))\, \rho^{\otimes N}\, (I - P_{N,\varepsilon}(\rho))}_{M}$ is positive so by cyclicity

$$\mathrm{Tr}\, S_N\, \rho^{\otimes N}(I - P_{N,\varepsilon}(\rho)) = \mathrm{Tr}\, M^{1/2}\, S_N\, M^{1/2}$$

$$\leq \mathrm{Tr}\, M^{1/2}\, I\, M^{1/2}$$

$$\leq \mathrm{Tr}\, M = \mathrm{Tr}\, \rho^{\otimes N}(I - P_{N,\varepsilon}(\rho))$$

$$\leq \delta.$$

Encoding - Decoding operations.

* Suppose that the input word is $|\varphi_{x_1}\rangle \otimes \cdots \otimes |\varphi_{x_N}\rangle$.
(this has a probability $p_{x_1} \cdots p_{x_N}$). We expect the encoding operation to compress this word to a string of M Qbits so an element of the Hilbert space $\mathbb{C}_2^{\otimes M} = \underbrace{\mathbb{C}_2 \otimes \cdots \otimes \mathbb{C}_2}_{M \text{ times}}$.

We allow for a slightly more general definition of the encoding map!

$$\mathcal{E} : \mathcal{H}^{\otimes N} \longrightarrow \{\text{Density matrices of } \mathbb{C}_2^{\otimes M}\}$$

$$|\varphi_{x_1}\rangle \otimes \cdots \otimes |\varphi_{x_N}\rangle \longmapsto \rho_M(x_1 \dots x_N).$$

where ρ_M is a $2^M \times 2^M$ density matrix. We require $M < N$ which corresponds to compression. The rate of the encoder is

$$R = \frac{M}{N} < 1.$$

* The decoding operation is a map:

$$\mathcal{D} : \{\text{Density matrices of } \mathbb{C}_2^{\otimes M}\} \longrightarrow \{\text{Density matrices of } \mathcal{H}^{\otimes N}\}$$

$$\rho_M(x_1 \dots x_N) \longmapsto \rho_N^{\text{output}}(x_1 \dots x_N).$$

Idealy we would like to recover the input word so we would like to have $\rho_N^{\text{output}}(x_1 \dots x_N) = (|\varphi_{x_1}\rangle \otimes \cdots \otimes |\varphi_{x_N}\rangle)(\langle\varphi_{x_1}| \otimes \cdots \otimes \langle\varphi_{x_N}|)$

The block diagram of the encoding-decoding scheme is

$$\text{source (Bob)} \longrightarrow \boxed{\text{Encoder}}_{\mathcal{E}} \longrightarrow \boxed{\begin{matrix}\text{Noiseless}\\ \text{channel or}\\ \text{storing medium}\end{matrix}} \longrightarrow \boxed{\text{Decoder}}_{\mathcal{D}} \longrightarrow \text{destination (Alice)}$$

$|\varphi_{x_1}\rangle \dots |\varphi_{x_N}\rangle$ $\qquad \rho_M \qquad \rho_N^{\text{output}}$.

$p_{x_1} \cdots p_{x_N}$

Reliability criterion.

Given that $|\varphi_{x_1} \dots \varphi_{x_N}\rangle$ was sent we define the "Fidelity" as

$$F(x_1 \dots x_N) = \langle \varphi_{x_1} \cdots \varphi_{x_N} | \rho_N^{\text{output}} | \varphi_{x_1} \cdots \varphi_{x_N} \rangle .$$

The motivation for this definition is that if the transmission is ideal we have $F = 1$. Moreover if the letters $\{|\varphi_x\rangle\}$ are all mutually orthogonal and $\rho_N^{\text{output}} = |\varphi_{y_1} \dots \varphi_{y_N}\rangle\langle \varphi_{y_1} \dots \varphi_{y_N}|$ then we are reduced to the classical case where

$$F(x_1 \dots x_N) = \begin{cases} 1 & \text{if } (x_1 \dots x_N) = (y_1 \dots y_N) \\ 0 & \text{if not} \end{cases}$$

The average fidelity is

$$F = \sum_{x_1 \dots x_N} p_{x_1} \cdots p_{x_N} \langle \varphi_{x_1} \cdots \varphi_{x_N} | \rho_N^{\text{output}} | \varphi_{x_1} \cdots \varphi_{x_N} \rangle .$$

In the classical case this reduces to the probability of block error.
Hereafter we take the average fidelity as our notion of reliability.
Note that one can make other definitions and it is not a priori

obvious that they are all equivalent. We are now ready to formulate our main theorem:

Theorem [Quantum Source Coding - Schumacher 1995].

Consider a source prepared with the ensemble $\{p_x, |\varphi_x\rangle\}$.

Take any $\varepsilon > 0$, $\delta > 0$ (small).

(A) Achievability part: let $R > S(\rho) + \varepsilon$. There exist an encoding-decoding scheme $(\mathcal{E}, \mathcal{D})$ with rate R such that for N sufficiently large $F \geq 1 - \delta$.

(B) Converse part: let $R < S(\rho) - \varepsilon$. For any scheme $(\mathcal{E}, \mathcal{D})$ with rate R we have for N sufficiently large that $F \leq \delta$.

Proof of Achievability:

As in the classical Shannon theorem the main idea is to encode and decode faithfully only the words that belong to the typical subspace. This will basically require $N S(\rho)$ Qbits. The rest of the words have small probability of occurrence so the probability of error will be small. The details are as follows:

Encoding operation:

Given an input word $|\varphi_{x_1} \dots \varphi_{x_N}\rangle$ do a measurement in the basis $(P_{N,\varepsilon}(\rho) \,;\, I - P_{N,\varepsilon}(\rho))$. The outcome is

$$\frac{P_{N,\varepsilon}(\rho)\,|\varphi_{x_1}\dots\varphi_{x_N}\rangle}{\langle\varphi_{x_1}\dots\varphi_{x_N}|\,P_{N,\varepsilon}(\rho)\,|\varphi_{x_1}\dots\varphi_{x_N}\rangle^{1/2}} \qquad \text{with prob} = \langle\varphi_{x_1}\dots\varphi_{x_N}|P_{N\varepsilon}(\rho)|\varphi_{x_1}\dots\varphi_{x_N}\rangle$$

or

$$\frac{(I - P_{N,\varepsilon}(\rho))\,|\varphi_{x_1}\rangle\dots|\varphi_{x_N}\rangle}{\langle\varphi_{x_1}\dots\varphi_{x_N}|\,I-P_{N\varepsilon}\,|\varphi_{x_1}\dots\varphi_{x_N}\rangle^{1/2}} \qquad \text{with prob} = \langle\varphi_{x_1}\dots\varphi_{x_N}|I-P_{N\varepsilon}(\rho)\,|\varphi_{x_1}\dots\varphi_{x_N}\rangle.$$

* If the outcome belongs to the typical subspace (first case) then the state $P_{N,\varepsilon}(\rho)\,|\varphi_{x_1}\dots\varphi_{x_N}\rangle$ has **at** most $2^{N(S(\rho)+\varepsilon)}$ non zero components. Note that the typical subspace vectors all have at most $2^{N(S(\rho)+\varepsilon)}$ non zero components. This means that by a transposition (or permutations) of components we can package all these vectors in a $2^{N(S(\rho)+\varepsilon)}$ dim vector (with the rest of the components being zero). This operation can be represented by a unitary matrix U:

$$U\,P_{N,\varepsilon}(\rho)\,|\varphi_{x_1}\dots\varphi_{x_N}\rangle = \sum C^{b_1\dots b_M}_{x_1\dots x_N}\,|\underbrace{b_1\dots b_M}_{\mathbb{C}_2^{\otimes M}},\underbrace{0,0\dots0}_{N-M}\rangle.$$

$$= |\psi^{\text{compressed}}_{x_1\dots x_N}\rangle \otimes |\underbrace{0,0\dots0}_{N-M}\rangle.$$

One sends $|\psi^{\text{compressed}}_{x_1\dots x_N}\rangle$ which is a M - qbit state.

* If the outcome belongs to the a-typical subspace (second case) the state $(I - P_{N\varepsilon}(\rho))|\varphi_{x_1} \dots \varphi_{x_N}\rangle$ is coded as a "junk state". In other words we send just $|junk\rangle$ which a specified state of the typical subspace, say

Thus at the output of the encoder, the density matrix is

$$\sigma = \langle \varphi_{x_1} \cdots \varphi_{x_N} | P_{N\varepsilon} | \varphi_{x_1} \dots \varphi_{x_N} \rangle \cdot | \psi^{\text{compressed}}_{x_1 \dots x_N} \rangle \langle \psi^{\text{compressed}}_{x_1 \dots x_N} |$$

$$+ \langle \varphi_{x_1} \cdots \varphi_{x_N} | I - P_{N\varepsilon} | \varphi_{x_1} \cdots \varphi_{x_N} \rangle \cdot |junk\rangle \langle junk|$$

To summarize the encoder maps

$$\mathcal{E} : |\varphi_{x_1} \cdots \varphi_{x_N}\rangle \langle \varphi_{x_1} \cdots \varphi_{x_N}| \longrightarrow \sigma .$$

Decoding Operation.

The input to the decoder is σ. First one appends the $(N-M)$ 0 states that were discarded before. Then one applies the inverse transposition or permutation (a unitary operation) to recover

$$\frac{P_{N,\varepsilon} |\varphi_{x_1} \cdots \varphi_{x_N}\rangle}{\langle \varphi_{x_1} \cdots \varphi_{x_N} | P_{N\varepsilon} | \varphi_{x_1} \cdots \varphi_N \rangle^{1/2}} .$$

So the decoder maps: $\mathcal{D} : \sigma \longrightarrow \rho^{\text{output}}$ with

$$\rho^{\text{output}} = P_{N,\varepsilon} |\varphi_{x_1} \cdots \varphi_{x_N}\rangle\langle \varphi_{x_1} \cdots \varphi_{x_N}| P_{N,\varepsilon}$$

$$+ \langle \varphi_{x_1} \cdots \varphi_{x_N} | (I - P_{N\varepsilon}) | \varphi_{x_1} \cdots \varphi_{x_N}\rangle U |\text{junk}, 0_{N-M}\rangle\langle \text{junk}, 0_{N-M}| U^{\dagger}$$

It remains to estimate the fidelity of this scheme.

$$F = \sum_{x_1 \cdots x_N} p_{x_1} \cdots p_{x_N} \langle \varphi_{x_1} \cdots \varphi_{x_N} | \rho^{\text{output}} | \varphi_{x_1} \cdots \varphi_{x_N}\rangle .$$

* The first contribution is

$$\sum_{x_1 \cdots x_N} p_{x_1} \cdots p_{x_N} |\langle \varphi_{x_1} \cdots \varphi_{x_N} | P_{N\varepsilon}(\rho) | \varphi_{x_1} \cdots \varphi_{x_N}\rangle|^2$$

$$\geq \left\{ \sum_{x_1 \cdots x_N} p_{x_1} \cdots p_{x_N} \langle \varphi_{x_1} \cdots \varphi_{x_N} | P_{N\varepsilon}(\rho) | \varphi_{x_1} \cdots \varphi_{x_N}\rangle \right\}^2 \quad \text{(Schwartz ineqn).}$$

$$= \left\{ \mathrm{Tr} \left(\rho^{\otimes N} P_{N,\varepsilon}(\rho) \right) \right\}^2 \geq (1-\delta)^2 \quad \text{(by thm on typical subspace).}$$

* The second contribution is

$$\sum_{x_1 \cdots x_N} p_{x_1} \cdots p_{x_N} \langle \varphi_{x_1} \cdots \varphi_{x_N} | I - P_{N\varepsilon}(\rho) | \varphi_{x_1} \cdots \varphi_{x_N}\rangle$$

$$\cdot |\langle \varphi_{x_1} \cdots \varphi_{x_N} | U | \text{junk}, 0_{N-M}\rangle|^2 \geq 0 .$$

This proves the achievability part.

Proof of the converse.

Let $\mathcal{E} : |\varphi_{x_1} \cdots \varphi_{x_N}\rangle\langle\varphi_{x_1} \cdots \varphi_{x_N}| \to \sigma$ be a general encoding scheme. Before decoding we append $|0_{N-M}\rangle\langle 0_{N-M}|$ as before so the input to the decoder is

$$\sigma \otimes |0_{N-M}\rangle\langle 0_{N-M}| .$$

This state belongs to a $2^M = 2^{NR}$ subspace of $\mathcal{H}^{\otimes N}$ with $R < S(\rho) - \varepsilon$. Let S_N be the projector on this subspace. We have

$$S_N \; \sigma \otimes |0_{N-M}\rangle\langle 0_{N-M}| \; S_N = \sigma \otimes |0_{N-M}\rangle\langle 0_{N-M}| .$$

Now we restrict ourselves to a unitary decoder $\mathcal{D}$ as before:

$$\mathcal{D} : \sigma \otimes |0_{N-M}\rangle\langle 0_{N-M}| \longrightarrow U \, \sigma \otimes |0_{N-M}\rangle\langle 0_{N-M}| \, U^+ .$$

So
$$\rho^{\text{output}} = U \, \sigma \otimes |0_{N-M}\rangle\langle 0_{N-M}| \, U^+$$
$$= U S_N \, \sigma \otimes |0_{N-M}\rangle\langle 0_{N-M}| \, S_N U^+ .$$

The Fidelity is

$$F = \sum_{x_1 \cdots x_N} p_{x_1} \cdots p_{x_N} \langle \varphi_{x_1} \cdots \varphi_{x_N} | U S_N \, \sigma \otimes |0_{N-M}\rangle\langle 0_{N-M}| \, S_N U^+ | \varphi_{x_1} \cdots \varphi_{x_N}\rangle$$

$$= \mathrm{Tr} \left(\rho^{\otimes N} U S_N \sigma \otimes |0_{N-M}\rangle\langle 0_{N-M}| S_N U^+ \right) .$$

$$= \mathrm{Tr} \left(S_N U^+ \rho^{\otimes N} U S_N \right) \left(\sigma \otimes |0_{N-M}\rangle\langle 0_{N-M}| \right) .$$

Since $S_N U^+ \rho^{\otimes N} U S_N$ is a positive operator and $\sigma \otimes |0_{N-M}\rangle\langle 0_{N-M}|$ is a density matrix and thus ≤ 1

we obtain $F \leq \mathrm{Tr}\ S_N U^+ \rho^{\otimes N} U$

$$= \mathrm{Tr}\ (U S_N U^+)\ \rho^{\otimes N} .$$

Now $U S_N U^+$ is a projector on a 2^{NR} subspace with $R < S(\rho) - \varepsilon$. Thus by the third statement of the typical subspace theorem we have $F \leq \delta$.

This achieves the proof of the converse for unitary decodings. For non-unitary decodings the proof is more difficult and is ommitted here. ∎

2.b) Source of mixed states.

For a source emitting mixed states ρ_x with probabilities p_x, the input messages are strings of N "letters"

$$\rho_{x_1} \otimes \cdots \otimes \rho_{x_N} .$$

In this case it is known that a rate $R < \chi(\rho) = S(\rho) - \sum p_x S(\rho_x)$ is not achievable (whatever the encoding-decoding scheme). However it is not known if a rate $R > \chi(\rho)$ is achievable although this is conjectured to be the case.

The proof that $R < \chi(\rho)$ is not achievable will be omitted in these notes.